HAL
open science

A Law of Large Numbers and Large Deviations for interacting diffusions on Erdős-Rényi graphs

Fabio Coppini, Helge Dietert, Giambattista Giacomin

To cite this version:

Fabio Coppini, Helge Dietert, Giambattista Giacomin. A Law of Large Numbers and Large Deviations for interacting diffusions on Erdős-Rényi graphs. Stochastics and Dynamics, 2020, 20 (2), pp.2050010. 10.1142/S0219493720500100 . hal-02145087

HAL Id: hal-02145087

https://hal.science/hal-02145087v1

Submitted on 1 Jun 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

arXiv:1807.10921v3 [math.PR] 29 May 2019

A LAW OF LARGE NUMBERS AND LARGE DEVIATIONS FOR INTERACTING DIFFUSIONS ON ERDŐS-RÉNYI GRAPHS

FABIO COPPINI, HELGE DIETERT, AND GIAMBATTISTA GIACOMIN

Abstract. We consider a class of particle systems described by differential equations (both stochastic and deterministic), in which the interaction network is determined by the realization of an Erdős-Rényi graph with parameter $p_n \in (0,1]$, where n is the size of the graph (i.e., the number of particles). If $p_n \equiv 1$ the graph is the complete graph (mean field model) and it is well known that, under suitable hypotheses, the empirical measure converges as $n \to \infty$ to the solution of a PDE: a McKean-Vlasov (or Fokker-Planck) equation in the stochastic case, or a Vlasov equation in the deterministic one. It has already been shown that this holds for rather general interaction networks, that include Erdős-Rényi graphs with $\lim_n p_n n = \infty$, and properly rescaling the interaction to account for the dilution introduced by p_n. However, these results have been proven under strong assumptions on the initial datum which has to be *chaotic*, i.e. a sequence of independent identically distributed random variables. The aim of our contribution is to present results – Law of Large Numbers and Large Deviation Principle – assuming only the convergence of the empirical measure of the initial condition.

2010 Mathematics Subject Classification: 60K35, 82C20

1. Introduction

1.1. **Basic notations, the models and a first look at the main question.** Large systems of interacting diffusions with mean field type interactions have been an important research topic in the mathematical community at least since the 60's. The program of identifying the emerging behavior for $n \to \infty$, where n is the number of interacting *units*, has been fully developed under suitable regularity and boundedness assumptions on the coefficients defining the system. In particular, Law of Large Numbers, Central Limit Theorems and Large Deviation Principles have been established (see for example [26, 21, 27, 7, 5]). A number of important issues remain unsolved, like the generalization to singular interactions (e.g. [16]) or understanding the delicate issue of considering at the same time large n and large time (e.g. [18]). But another direction in which mathematical results are still very limited is about relaxing the *complete graph assumption* for the interaction network – *complete graph* is just a different wording for *mean field* – and going towards more heterogeneous interaction networks. This is an issue that emerges in plenty of applied disciplines and giving a proper account of the available literature would be a daunting task: so we limit ourselves to signaling the recent survey [25] which contains an extended literature.

We are therefore going to study the emerging behavior of interacting diffusion models when, like in complete graphs, every unit interacts with a diverging number of other units. The interaction network is described as a random graph, notably of Erdős-Rényi (ER) type; so we start with the basic notions on graphs.

Let $\xi^{(n)} = \{\xi^{(n)}_{i,j}\}_{i,j\in\{1,\dots,n\}}$ denote the adjacency matrix of a graph $\left(V^{(n)}, E^{(n)}\right)$ with n vertices ($\xi^{(n)}$ will also denote the graph itself):

$$V^{(n)} := \{1,\dots,n\} \quad \text{and} \quad E^{(n)} := \left\{(i,j) \in V^{(n)} \times V^{(n)} : \xi^{(n)}_{i,j} = 1\right\} . \tag{1.1}$$

We consider sequences of asymmetric ER random graphs with self loops with probabilities $p_n \in (0,1)$ for $n = 2,3,\dots$. More precisely, we just assume that $\{\xi^{(n)}_{i,j}\}_{i,j\in\{1,\dots,n\}}$ are Independent Identically Distributed (IID) Bernoulli random variables of parameter p_n (with notation $\mathrm{B}(p_n)$). The arguments are easily adapted to the case in which $\xi^{(n)}_{j,j} = 0$ for every j and the results are unchanged.

Even if these graphs are not coupled for different values of n, it is practical to work with only one probability space and to couple these adjacency matrices (or random graphs). For example one can start from a sequence $\{U_k\}_{k\in\mathbb{N}}$ of IID $U(0,1)$ variables and define $\xi^{(n)}_{i,j} = \mathbf{1}_{U_{k(i,j)}<p_n}$, with k an arbitrary bijection from $\mathbb{N}^2$ to $\mathbb{N}$. The law of the graph is denoted by $\mathbb{P}$, with $\mathbb{E}$ the corresponding expectation, and we will just write $\mathbb{P}(\,\mathrm{d}\xi)$-a.s. meaning "almost surely in the realization of $\{\xi^{(n)}\}_{n=2,3,\dots}$".

Given a realization of $\xi^{(n)}$, consider the n-dimensional diffusion $\theta^n_t := \{\theta^{i,n}_t\}_{i=1,\dots,n}$ which solves for every i

$$\mathrm{d}\theta^{i,n}_t = F\left(\theta^{i,n}_t\right)\mathrm{d}t + \frac{1}{n}\sum_{j=1}^{n}\frac{\xi^{(n)}_{i,j}}{p_n}\Gamma\left(\theta^{i,n}_t, \theta^{j,n}_t\right)\mathrm{d}t + \sigma\left(\theta^{i,n}_t\right)\mathrm{d}B^i_t, \tag{1.2}$$

where $\{B^i_\cdot\}_{i\in\mathbb{N}}$ are independent standard Brownian motions (whose law is denoted by $\mathbf{P}$) and independent also of $\xi^{(n)}$ (so, we are effectively working with $\mathbf{P}\otimes\mathbb{P}$). For simplicity, we consider only deterministic initial conditions; but the results apply to random initial conditions once they are taken independent of Brownian motions and of ξ. Moreover, assume that:

(1) F, Γ and σ are real valued (uniformly) Lipschitz functions: the corresponding Lipschitz constants are denoted by L_F, L_Γ and L_σ;
(2) Γ is bounded, in particular $\|\Gamma\|_\infty := \sup_{x,y\in\mathbb{R}}|\Gamma(x,y)| < \infty$;
(3) $\sigma_- \le \sigma(\cdot) \le \sigma_+$ with $\sigma_\pm$ two positive constants (non degenerate diffusion). If $\sigma(\cdot)$ is a constant, we include the case $\sigma(\cdot)\equiv 0$.

Fix $T > 0$, the law of the n trajectories $\{\theta^n_t\}_{t\in[0,T]}$ for the *quenched* system is denoted by $\mathbf{P}^\xi_n$, i.e. $\mathbf{P}^\xi_n \in \mathcal{P}\left(\mathcal{C}^0([0,T];\mathbb{R}^n)\right)$, and the associated empirical measure at time t by $\{\mu^n_t\}_{t\in[0,T]}$, i.e.

$$\mu^n_t := \frac{1}{n}\sum_{j=1}^{n}\delta_{\theta^{j,n}_t} \in \mathcal{P}(\mathbb{R}) . \tag{1.3}$$

$\mathcal{P}(\mathbb{R})$ denotes the set of probability measures over $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ equipped with the (metrizable) topology of weak convergence: i.e., if $\mu_n \in \mathcal{P}(\mathbb{R})$ for every n, then $\lim_n \mu_n = \mu \in \mathcal{P}(\mathbb{R})$ if $\int h(x)\mu_n(\,\mathrm{d}x) \to \int h(x)\mu(\,\mathrm{d}x)$ as $n\uparrow\infty$ for every $h(\cdot)$ continuous and bounded function. Note that since ξ is random, μ^n_t is a random variable taking values in $\mathcal{P}(\mathbb{R})$, equipped with the σ-algebra of its Borel subsets.

The solution $\{\theta_t^{i,n}\}_{i=1,\dots,n}$ is going to be tightly linked with $\{\bar\theta_t^{i,n}\}_{i=1,\dots,n}$ which solves

$$\mathrm{d}\bar\theta_t^{i,n} \,=\, F\left(\bar\theta_t^{i,n}\right)\,\mathrm{d}t + \frac{1}{n}\sum_{j=1}^n \Gamma\left(\bar\theta_t^{i,n},\bar\theta_t^{j,n}\right)\,\mathrm{d}t + \sigma\left(\bar\theta_t^{i,n}\right)\,\mathrm{d}B_t^i. \tag{1.4}$$

The law of $\{\bar\theta_t^n\}_{t\in[0,T]}$ is denoted by $\mathbf{P}_n$. Moreover $\bar\mu_t^n := (1/n)\sum_{j=1}^n \delta_{\bar\theta_t^{j,n}}$. Often (1.4) is called *annealed* system: of course (1.4) is obtained from (1.2) by taking the expectation of the drift with respect to $\mathbb{P}$.

If the empirical measure of the initial conditions converges to a probability μ_0, i.e.

$$\lim_{n\to\infty} \bar\mu_0^n \,=\, \mu_0 \in \mathcal{P}\left(\mathbb{R}\right)\,, \tag{1.5}$$

and if $\int_{\mathbb{R}} x^2\mu_0(\,\mathrm{d}x) < \infty$, then it is well known that, for every $t>0$, $\bar\mu_t^n$ weakly converges in $\mathcal{P}\left(\mathbb{R}\right)$ to μ_t, the unique weak solution of the following McKean-Vlasov (or Fokker-Planck) equation

$$\partial_t\mu_t(\theta) \,=\, \frac{1}{2}\partial_\theta^2\left(\sigma^2(\theta)\mu_t(\theta)\right) - \partial_\theta\left(\mu_t(\theta)F(\theta)\right) - \partial_\theta\left(\mu_t(\theta)\int_{\mathbb{R}}\Gamma(\theta,\theta')\mu_t(\,\mathrm{d}\theta')\right)\,. \tag{1.6}$$

The slightly stronger result that is proven is in fact: for every $T>0$, if one considers $\mu_\cdot^n$ as an element of $C^0([0,T];\mathcal{P}(\mathbb{R}))$ (a complete separable metric space), then $\lim_n \mu_\cdot^n = \mu_\cdot$ ($\mathbf{P}$-a.s. when σ is non degenerate). The notion of weak solution $\mu_\cdot \in C^0([0,T];\mathcal{P}(\mathbb{R}))$ to (1.6), which can be found for example in [14], is strictly related to the nonlinear diffusion formulation: the stochastic process $\{\varphi_t\}_{t\in[0,T]}$ that solves

$$\begin{cases} \mathrm{d}\varphi_t \,=\, F\left(\varphi_t\right)\,\mathrm{d}t + \int\Gamma\left(\varphi_t,\varphi\right)\nu_t\left(\,\mathrm{d}\varphi\right)\,\mathrm{d}t + \sigma\left(\varphi_t\right)\,\mathrm{d}B_t\,, \\ \nu_t \,=\, \mathrm{Law}(\varphi_t)\,, \quad \text{for all } t\in[0,T]\,, \end{cases} \tag{1.7}$$

with initial condition which is a square integrable random variable independent of the standard Brownian motion $B_\cdot$. Existence and uniqueness for this *atypical* stochastic differential equation is not obvious at all, but it is by now well known that if $\nu_0=\mu_0$, then the unique $\nu_\cdot \in C^0([0,T];\mathcal{P}(\mathbb{R}))$ such that ν_t is the law of φ_t for all $t\in[0,T]$, is the unique weak solution of (1.6), i.e. $\nu_t=\mu_t$ for all $t\in[0,T]$. The literature on the results that we have just mentioned is vast, see e.g. [23, 26, 21, 14] for the non degenerate diffusion case and [10, 22] for the $\sigma(\cdot)\equiv 0$ case; in this last case there is no need to assume that $\int_{\mathbb{R}} x^2\mu_0(\,\mathrm{d}x)<\infty$.

In the sequel, we will also work with probabilities in $\mathcal{P}\left(\mathcal{C}^0([0,T];\mathbb{R})\right)$, that is considering the law of $\{\varphi_t\}_{t\in[0,T]}$ seen as a random trajectory on the path space $\mathcal{C}^0([0,T];\mathbb{R})$, rather than its time marginals $\mu_t\in\mathcal{P}(\mathbb{R})$.

Remark 1.1. *Observe that knowing the law of* (1.7) *gives more information than the solution $\mu_\cdot$ of the McKean-Vlasov equation* (1.6)*. Indeed, call P_φ the law of $\{\varphi_t\}_{t\in[0,T]}$, then P_φ is an element of $\mathcal{P}\left(\mathcal{C}([0,T];\mathbb{R})\right)$, whereas $\mu_\cdot\in C^0([0,T];\mathcal{P}(\mathbb{R}))$. It is straighforward to obtain μ_t from P_φ by just observing*

$$\mu_t(\cdot) = P_\varphi\circ\pi_t^{-1}(\cdot), \tag{1.8}$$

where $\pi_t : \mathcal{C}([0,T];\mathbb{R})\to\mathbb{R}$ is the canonical projection at time t. Observe that a reverse statement is not always possible: $\mu_\cdot$ alone does not allow to compute multidimensional time marginals like $\mathbf{P}(\varphi_s\in A,\varphi_t\in B)$, for $s,t\in[0,T]$ and $A,B\subset\mathbb{R}$. Existence, uniqueness and well-posedness of the problem for P_φ can be found in [21] *and references therein.*

1.2. **Aim of the paper.** Informally stated, our aim is to study the proximity of $\mu^n_\cdot$ and $\bar{\mu}^n_\cdot$, for n large. Since $\bar{\mu}^n_\cdot$ approaches the solution of the McKean-Vlasov equation (1.6), this turns out to be studying the proximity of $\mu^n_\cdot$ and the solution of the McKean-Vlasov equation. This of course requires (at least) the assumption that

$$\lim_{n\to\infty} \mu^n_0 \,=\, \mu_0\,. \tag{1.9}$$

A result of this type has been already achieved: in the case $\sigma(\cdot) \equiv \sigma \geq 0$, [8] proved a LLN for the trajectories of (1.2) where $\xi^{(n)}$ is a (deterministic) sequence of graphs such that

$$\lim_{n\to\infty} \sup_{i\in\{1,\dots,n\}} \left| \frac{1}{n} \sum_{j=1}^{n} \frac{\xi^{(n)}_{i,j}}{p_n} - 1 \right| \,=\, 0\,, \tag{1.10}$$

and with IID initial conditions (*chaotic* initial datum), that is $\theta^{j,n}_0 = \theta^j_0$ for every n and every $j = 1, \dots, n$ where

$$\left\{\theta^j_0\right\}_{j\in\mathbb{N}} \text{ is a typical realization of an IID sequence of variables with law } \mu_0\,. \tag{1.11}$$

Under conditions (1.10) and (1.11), it is proved that $\lim_n \mu^n_\cdot = \mu_\cdot$ in $\mathbf{P}$-probability. We recall that, as stated right after (1.2), $\{\theta^j_0\}_{j\in\mathbb{N}}$ is independent of the driving Brownians and of the graph ξ.

This seems at first rather satisfactory. However in [8] it is discussed at length how this result in reality is, on one hand, surprising and, on the other, that it does not really solve the problem. This can be understood by considering that the *homogeneous degree condition* (1.10) is $\mathbb{P}\,(\,\mathrm{d}\xi)$-a.s. verified for ER type graphs when $\liminf_n np_n/\log n$ is larger than a well-chosen constant (see [8, Proposition 1.3]). But the class of graphs satisfying (1.10) goes well beyond ER graphs: in particular, it is straightforward to construct graphs with an arbitrary number of connected components that satisfy (1.10), see the following remark.

Remark 1.2. (1.9) *and* (1.10) *are not sufficient to obtain a result in the direction we are aiming at. In fact, if* $\xi^{(n)}$ *is the graph in which two vertices are connected if and only if they have the same parity (which corresponds to* $\lim_n p_n = 1/2$*), then, as long as* μ_0 *is not the uniform measure, one can easily arrange the initial condition in order to have different limit distributions on even and odd sites, or no limit at all. Thus, as* $n \to \infty$*, the evolution will not be described by* (1.6)*.*

In a nutshell, the results in [8] are obtained under a weak assumption on the graph, but under strong assumptions on the initial condition. And this to the point of obtaining a result that is troublesome: a system with plenty of disconnected components behaves essentially like a totally connected one! Of course the *solution* of this apparent paradox is in the chaotic character of the initial condition that leads to a homogeneous and identical behavior of the initial datum on all components, and the fact that chaos propagates at least on a finite time horizon (see [8] for more on this issue). But there is no reason to expect mean field type behavior, assuming only (1.10) on the graph, without a strong *statistical homogeneity* assumption on the initial datum, as argued in Remark 1.2.

The aim of this paper is to attack the problem assuming only the convergence of the empirical measure of the initial datum, that is (1.9), but assuming that the graph is of ER type. Otherwise said, we want to make a minimal assumption on the initial condition and we try to exploity the *chaoticity* of the graph to achieve the result. We will attack

the problem from more then one perspective, not only the direct LLN angle of attack, but also from the Large Deviations (LD) perspective. The vast literature related to our results is presented and discussed after the statements.

2. Main results

Let us denote $d_{\mathrm{bL}}(\cdot,\cdot)$ the bounded Lipschitz distance which endows the weak convergence topology on $\mathcal{P}(\mathbb{R})$ (this choice is somewhat arbitrary: other distances can be used, for example the Wasserstein one, see [10]). By this we mean that $d_{\mathrm{bL}}(\mu,\nu) = \sup_h \left|\int h\,\mathrm{d}\mu - \int h\,\mathrm{d}\nu\right|$, where the supremum is taken over $h:\mathbb{R}\to[0,1]$ such that $|h(x)-h(y)|\le|x-y|$.

We are now ready to state the LLN. Recall that $\mu^n_\cdot$ is a random element of $C^0([0,T];\mathcal{P}(\mathbb{R}))$ and that $\mu_\cdot$, a non random element of $C^0([0,T];\mathcal{P}(\mathbb{R}))$, is the unique weak solution of the McKean-Vlasov equation (1.6).

Theorem 2.1. *Assume that the initial datum is deterministic, that it satisfies* (1.9) *and, if $\sigma(\cdot)\not\equiv 0$, that it satisfies also that $\int_{\mathbb{R}} x^2\mu_0(\,\mathrm{d}x)<\infty$. Make the hypothesis that p_n satisfies*

$$\liminf_{n\to\infty}\frac{p_n n}{\log n}>0\,, \tag{2.1}$$

and either that $0<\sigma_-\le\sigma(\cdot)\le\sigma_+<\infty$ or $\sigma(\cdot)\equiv 0$. Then $\mathbf{P}\otimes\mathbb{P}$-a.s. we have that

$$\lim_{n\to\infty}\mu^n_\cdot=\mu_\cdot \quad \text{in } C^0([0,T];\mathcal{P}(\mathbb{R})). \tag{2.2}$$

The requirement of deterministic initial data is easily lifted to IID initial conditions under the assumption that they are independent of the graph (and, of course, of the driving Brownians).

From the viewpoint of the proof, Theorem 2.1 may be viewed as two different statements.

- in the case of $\sigma(\cdot)\equiv\sigma\in[0,\infty)$, the proof follows by coupling the system on the ER graph and the system on the complete graph;
- in the case of $0<\sigma_-\le\sigma(\cdot)\le\sigma_+<\infty$, the result is a corollary of a Large Deviation Principle (LDP) stating that, at the Large Deviations (LD) level, the system on ER graph and the complete graph system are indistinguishable, see Theorem 2.2.

In the next subsection we present the result related to Large Deviations.

2.1. The Large Deviation Principle.

Stating the LDP needs some preparation on the general LD approach (classical references are for example [9, 11, 13]).

Given a complete, separable metric space χ, a rate function I is a lower semicontinuous mapping $I:\chi\to[0,\infty]$ such that each level set $K_l=\{x\in\chi: I(x)\le l\}$ is compact for all $l\ge 0$ (sometimes I is called a *good* rate function). Given $\{P_n\}_{n\in\mathbb{N}}$ a sequence of probability measures on χ associated with its Borel σ-field, we say that P_n satisfies a LDP (on χ) with rate function I if for every measurable set $A\subset\chi$

$$-\inf_{x\in A^\circ} I(x)\le\liminf_{n\to\infty}\tfrac{1}{n}\log P_n\left(A^\circ\right)\le\limsup_{n\to\infty}\tfrac{1}{n}\log P_n\left(\bar{A}\right)\le-\inf_{x\in\bar{A}} I(x), \tag{2.3}$$

where A° is the interior of A and $\bar{A}$ is its closure.

Let us now recall that (1.4), or equivalently (1.2) on a complete graph, satisfies a LDP, we refer to [5, Theorem 3.1]. We choose to state the LDP for the empirical law of the process, that is for

$$\overline{L}_n := \frac{1}{n}\sum_{j=1}^{n}\delta_{\bar{\theta}^{j,n}_{\cdot}} \in \mathcal{P}\left(C^0([0,T];\mathbb{R})\right), \tag{2.4}$$

but other LDP are possible. Namely, [7, Theorem 5.1] proves a LDP for the empirical measure $\bar{\mu}^n_{\cdot}$ seen as an element of $C^0\left([0,T];\mathcal{P}(\mathbb{R})\right)$ (recall (1.8)), yet our result includes this case. In Remark 1.1 we have pointed out the continuity of the projection π_t and how to pass from P_n to $\mu^n_{\cdot} \in C^0([0,T];\mathcal{P}(\mathbb{R}))$, therefore a corollary of a LDP for $\bar{L}_n$ is a LDP on $C^0([0,T];\mathcal{P}(\mathbb{R}))$ for the law of $\bar{\mu}^n_{\cdot}$ with LD functional given by the contraction principle: see for example [4, 17] for the mathematical procedure and [7] for an explicit form of the LD functional in the full generality.

We set $\chi = \mathcal{P}\left(C^0([0,T];\mathbb{R})\right)$; since $C^0([0,T];\mathbb{R})$ is a metric space, χ is a complete, separable metric space once equipped (among various possibilities) with the bounded Lipschitz distance. Define the probability measure $\overline{P}_n$ on χ by setting $\overline{P}_n(\cdot) := \mathbf{P}(\overline{L}_n \in \cdot)$, of course χ equipped with the σ-algebra of its Borel subsets, then [5, Theorem 3.1] shows that $\overline{P}_n$ satisfies a LDP whose rate function concentrates on $\nu \in \chi$ such that $\nu_{\cdot} = \nu \circ \pi_{\cdot}^{-1} \in C^0([0,T];\mathcal{P}(\mathbb{R}))$ is solution of the McKean-Vlasov equation (1.6).

We are now ready to state the main result of this subsection. For every realization of the graph ξ define the probability P_n^{ξ} on χ, by setting $P_n^{\xi}(\cdot) := \mathbf{P}(L^n \in \cdot)$, where $L^n(\cdot)$ is defined as in (2.4), but replacing $\bar{\theta}^{j,n}_{\cdot}$ with $\theta^{j,n}_{\cdot}$. In particular, P_n^{ξ} is the empirical measure of the trajectories $\theta^{j,n}_{\cdot}$ solving (1.2).

Theorem 2.2. *Assume that $\sigma_- > 0$. If ξ is an ER graph that satisfies* (2.1) *and if the initial datum satisfies* (1.9) *and $\int x^2\mu_0(\mathrm{d}x) < \infty$, then P_n^{ξ} satisfies the same LDP of P_n $\mathbb{P}(\mathrm{d}\xi)$-a.s..*

2.2. **A look at the literature.** We recall that for interacting particle systems on the complete graph, i.e. (1.4), many results on the LLN are available and many of them, as [8], include propagation of chaos properties. However, as already mentioned, propagation of chaos results are very demanding on the initial condition.

The literature is vast and difficult to be properly cited: we mention the seminal contribution [19] and we mention again [23, 26, 21, 14], that are also useful source of more references and that are not limited to propagation of chaos results, in the sense that also the case of deterministic initial data is treated. For the $\sigma(\cdot) \equiv 0$ case, we mention the important original contributions [10, 22] that gave origin to a vast literature that goes beyond our purposes.

Large deviation properties for mean field diffusions have been studied in the seminal work by Dawson and Gärtner [7], but also in [13, 12, 17] in the so called gradient case. In [5] the problem is attacked in great generality using an approach based on weak convergence and control theory.

The LLN case has already been adressed in the literature, even if few results seem to have been proven so far. As mentioned, [8] proves a LLN for $\mu^n_{\cdot}$ requiring the initial datum to be a product measure: the case $\sigma(\cdot) \equiv \sigma \geq 0$ is considered. In the same spirit, from the initial datum viewpoint, but for a time-varying graph and for multi-type processes, there is the work of [1]. It is important to mention at this stage that in [1] the interaction

is renormalized by the number of neighbors of each site i: we normalize instead by the expected number of neighbors.

Turning to LD results, the recent work of [24] extends the LDP for Hamiltonian systems in random media, presented in [4], to (sparse) random interactions which include symmetric ER random graphs. The convergence of the empirical measure is shown under the assumption $\lim_n np_n = \infty$, without requiring any log divergence. However, they still focus on IID initial conditions and constant diffusion term $\sigma(\cdot) \equiv 1$.

Focusing on the case $\sigma(\cdot) \equiv 0$, we mention the contributions:

- in [3] one finds the stability analysis for the stationary state of an ordinary differential equations system with ER interacting network, requiring a logarithmic divergence of $p_n n$;
- in [6] the Kuramoto model, i.e. $\Gamma(x, y) = \sin(x - y)$ and $F(\cdot)$ is a random constant (*natural frequencies*), is studied with an interaction network that is given by a *graphon*: this leads to a more general limit equation, but their approach includes the case of ER graphs (in this case the graphon is trivial) with p_n that tends to a positive constant. In [20] the case of sparse graphs is considered: for ER graphs the condition is $\lim_n p_n \sqrt{n} = \infty$.

In many of the papers we cite, notably [8, 4, 17, 6, 20], another source of randomness is allowed: for example, in the Kuramoto model this corresponds to the important feature that each oscillator has a priori its own oscillation frequency and, more generally, with this extra source of randomness we can model systems in which the interacting diffusions (or units, agents,...) are not identical. This source of randomness is chosen independently of the graph and of the dynamical noise. All the results we have presented generalize easily to this case, but at the expense of heavier notations and heavier expressions. We have chosen not to treat this case for sake of conciseness and readability.

The rest of the paper is devoted to the proofs. Section 3 contains the proof of Theorem 2.1 in the case of constant (possibly degenerate) diffusion coefficient. Section 4 contains the proof of Theorem 2.2.

3. The Law of Large Numbers: the proof

This section is devoted to the proof of Theorem 2.1 in the case $\sigma(\cdot) \equiv \sigma \in [0, \infty)$: we recall that the case of non trivial and non degenerate diffusion is a corollary of the LDP (Theorem 2.2). We start with two preliminary lemmas that will be used for Proposition 3.3, from which Theorem 2.1 follows.

Lemma 3.1. *Let $K > 2$. For all $n \in \mathbb{N}$, it holds*

$$\mathbb{P}\left(\sum_{j=1}^{n} \left|\frac{\xi_{i,j}^{(n)}}{p_n} - 1\right| \geq Kn\right) \leq \exp\left(-\frac{3(K-2)^2}{6 + 2(K-2)} p_n n\right). \tag{3.1}$$

In particular, under hypothesis (2.1) and setting $C := \liminf_{n\to\infty} \frac{p_n n}{\log n} \in (0, \infty]$, we have that, if $K > K_C := 2 + \frac{2}{3C} + \sqrt{\frac{4}{9C^2} + \frac{4}{C}}$, then $\mathbb{P}(\mathrm{d}\xi)$-a.s. there exists $n_0 = n_0(\xi) < \infty$ such that for $n \geq n_0$

$$\max\left(\sup_{i=1,\dots,n} \sum_{j=1}^{n} \left|\frac{\xi_{j,i}^{(n)}}{p_n} - 1\right|, \sup_{i=1,\dots,n} \sum_{j=1}^{n} \left|\frac{\xi_{i,j}^{(n)}}{p_n} - 1\right|\right) \leq Kn. \tag{3.2}$$

Proof. We use Bernstein's inequality (see for example [2, Corollary 2.11]) which says that if $X_1, \dots, X_n$ are independent zero-mean random variables such that $|X_j| \le M$ a.s. for all j, then for all $t \ge 0$

$$\mathbb{P}\left(\sum_{j=1}^n X_j > t\right) \le \exp\left\{-\frac{\frac{1}{2}t^2}{\sum_{j=1}^n \mathbb{E}[X_j^2] + \frac{1}{3}Mt}\right\}.$$

Set $X_j = \left|\frac{\xi_{i,j}^{(n)}}{p_n} - 1\right| - 2(1-p_n)$. X_j is a zero-mean random variable and we can bound

$$\mathbb{P}\left(\sum_{j=1}^n \left|\frac{\xi_{i,j}^{(n)}}{p_n} - 1\right| \ge Kn\right) \le \mathbb{P}\left(\sum_{j=1}^n X_j \ge (K-2)n\right). \tag{3.3}$$

We have $|X_j| \le \max(1/p_n - 3 + 2p_n, 2p_n - 1) \le 1/p_n =: M$ and $\mathbb{E}[X_j^2] - 1/p_n = -5 + 8p_n - 4p_n^2 \le -1$, so $\mathbb{E}[X_j^2] \le 1/p_n$ and Bernstein's inequality together with an union bound show that

$$\mathbb{P}\left(\sup_{i=1,\dots,n} \sum_{j=1}^n \left|\frac{\xi_{i,j}^{(n)}}{p_n} - 1\right| \ge Kn\right) \le n \exp\left(-\frac{3(K-2)^2}{6+2(K-2)} p_n n\right). \tag{3.4}$$

The proof is now completed with some elementary computations and by applying the Borel-Cantelli Lemma. □

Lemma 3.2. *Assume* (2.1) *and let*

$$\Delta_i(s) := \left|\frac{1}{n}\sum_{j=1}^n \left(\frac{\xi_{i,j}^{(n)}}{p_n} - 1\right) \Gamma\left(\bar\theta_s^{i,n}, \bar\theta_s^{j,n}\right)\right|^2, \quad \textit{for every } s \in [0,T]. \tag{3.5}$$

Then, for every realization of the Brownian motions, it holds that

$$\lim_{n\to\infty} \int_0^T \frac{1}{n}\sum_{i=1}^n \Delta_i(s)\,\mathrm{d}s = 0, \quad \mathbb{P}(\,\mathrm{d}\xi)\text{-}a.s.. \tag{3.6}$$

Proof. First, we rewrite $\int_0^T \Delta_i(s)\,\mathrm{d}s$ as

$$\int_0^T \Delta_i(s)\,\mathrm{d}s = \frac{1}{(np_n)^2}\sum_{j,k=1}^n \hat\xi_{i,j}\hat\xi_{i,k} d_{ijk}, \tag{3.7}$$

where we have dropped the superscript (n), the dependency on T and we have introduced the notations

$$\hat\xi_{i,j} := \xi_{i,j}^{(n)} - p_n \quad \text{and} \quad d_{ijk} := \int_0^T \left[\Gamma\left(\bar\theta_s^{i,n}, \bar\theta_s^{j,n}\right)\Gamma\left(\bar\theta_s^{i,n}, \bar\theta_s^{k,n}\right)\right]\mathrm{d}s\,. \tag{3.8}$$

Observe that $\hat\xi_{i,j}$ are centered random variables and $|d_{ijk}| \le T\,\|\Gamma\|_\infty^2 =: d_\star$.

Let δ_n be a sequence of positive numbers such that (recall (2.1))

$$\delta_n \gg \frac{1}{p_n n} \quad \text{and} \quad \lim_{n\to\infty} \delta_n = 0. \tag{3.9}$$

Let Ω_n be the set

$$\Omega_n := \left\{ \xi : \frac{1}{(np_n)^2} \sum_{i,j,k=1}^{n} \hat{\xi}_{i,j}\hat{\xi}_{i,k} d_{ijk} > \delta_n n \right\}, \tag{3.10}$$

We want to show that

$$\sum_{n\in\mathbb{N}} \mathbb{P}\left(\Omega_n\right) < \infty. \tag{3.11}$$

Let $K > 2$ and consider the events

$$A_n = \bigcup_{i=1}^{n} A_{n,i} \quad \text{with} \quad A_{n,i} = \left\{ \xi^{(n)} : \sum_j \left| \frac{\xi_{i,j}^{(n)}}{p_n} - 1 \right| > Kn. \right\}. \tag{3.12}$$

We use

$$\mathbb{P}\left(\Omega_n\right) \leq \mathbb{P}\left(\Omega_n \cap A_n^{\complement}\right) + \mathbb{P}\left(A_n\right). \tag{3.13}$$

and Lemma 3.1 ensures that choosing $K > K_C (> 2)$ we have

$$\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) < \infty, \tag{3.14}$$

so that one is left with proving that $\sum_{n\in\mathbb{N}} \mathbb{P}\left(\Omega_n \cap A_n^{\complement}\right) < \infty$. By Markov's inequality applied to $\mathbb{P}\left(\cdot | A_n^{\complement}\right)$ we see that

$$\mathbb{P}(\Omega_n \cap A_n^{\complement}) \leq \exp\left(-n\delta_n + \log \mathbb{E}\left[\mathbf{1}_{A_n^{\complement}} \exp\left(\frac{1}{(np_n)^2} \sum_{i,j,k=1}^{n} \hat{\xi}_{i,j}\hat{\xi}_{i,k} d_{ijk} \right) \right] \right). \tag{3.15}$$

Given (3.9), it suffices to show that

$$\log \mathbb{E}\left[\mathbf{1}_{A_n^{\complement}} \exp\left(\frac{1}{(np_n)^2} \sum_{i,j,k=1}^{n} \hat{\xi}_{i,j}\hat{\xi}_{i,k} d_{ijk} \right) \right] = n\, O\left(\frac{1}{p_n n}\right) = O\left(\frac{1}{p_n}\right). \tag{3.16}$$

We exploit the independence w.r.t. i:

$$\mathbb{E}\left[\mathbf{1}_{A_n^{\complement}} \exp\left(\frac{1}{(np_n)^2} \sum_{i,j,k=1}^{n} \hat{\xi}_{i,j}\hat{\xi}_{i,k} d_{ijk} \right) \right] = \prod_i \mathbb{E}\left[\mathbf{1}_{A_{n,i}^{\complement}} \exp\left(\frac{1}{(np_n)^2} \sum_{j,k=1}^{n} \hat{\xi}_{i,j}\hat{\xi}_{i,k} d_{ijk} \right) \right], \tag{3.17}$$

and use the inequality $\exp(x) \le 1 + |x| \exp|x|$ which holds for all $x \in \mathbb{R}$, together with Cauchy-Schwarz and obtain

$$\begin{aligned}
&\mathbb{E}\left[\mathbf{1}_{A^{\complement}_{n,i}} \exp\left(\frac{1}{(np_n)^2}\sum_{j,k}\hat{\xi}_{i,j}\hat{\xi}_{i,k}d_{ijk}\right)\right] \le \\
&1 + \mathbb{E}\left[\mathbf{1}_{A^{\complement}_{n,i}}\left|\frac{1}{(np_n)^2}\sum_{j,k}\hat{\xi}_{i,j}\hat{\xi}_{i,k}d_{ijk}\right| \exp\left(\left|\frac{1}{(np_n)^2}\sum_{j,k}\hat{\xi}_{i,j}\hat{\xi}_{i,k}d_{ijk}\right|\right)\right] \le \\
&1 + \mathbb{E}\left[\left(\frac{1}{(np_n)^2}\sum_{j,k}\hat{\xi}_{i,j}\hat{\xi}_{i,k}d_{ijk}\right)^2\right]^{1/2} \mathbb{E}\left[\mathbf{1}_{A^{\complement}_{n,i}} \exp\left(\frac{2}{(np_n)^2}\sum_{j,k}\hat{\xi}_{i,j}\hat{\xi}_{i,k}d_{ijk}\right)\right]^{1/2}.
\end{aligned} \tag{3.18}$$

Under the condition that we are in $A^{\complement}_{n,i}$, it holds that

$$\left|\frac{2}{(np_n)^2}\sum_{j,k}\hat{\xi}_{i,j}\hat{\xi}_{i,k}d_{ijk}\right| \le 2K^2 d_\star \tag{3.19}$$

so that the exponential expectation can be bounded as $\exp\left\{2K^2 d_\star\right\}$. Estimating the moment expectation leads to

$$\mathbb{E}\left[\left(\frac{1}{(np_n)^2}\sum_{j,k}\hat{\xi}_{i,j}\hat{\xi}_{i,k}d_{ijk}\right)^2\right] = \frac{1}{(np_n)^4}\sum_{j,k,p,q}\mathbb{E}\left[\hat{\xi}_{i,j}\hat{\xi}_{i,k}\hat{\xi}_{i,p}\hat{\xi}_{i,q}\right]d_{ijk}d_{ipq} \le \tag{3.20}$$

$$\le \frac{d_\star^2}{(np_n)^4}\left[np_n + 3(np_n)^2\right] \le \frac{4d_\star^2}{(np_n)^2}. \tag{3.21}$$

From (3.18), we get

$$\mathbb{E}\left[\mathbf{1}_{A^{\complement}_{n,i}} \exp\left(\frac{1}{(np_n)^2}\sum_{j,k}\hat{\xi}_{i,j}\hat{\xi}_{i,k}d_{ijk}\right)\right] \le 1 + \frac{2d_\star}{np_n}\exp\left\{2K^2 d_\star\right\}. \tag{3.22}$$

Putting everything back in (3.17), one obtains

$$\mathbb{E}\left[\mathbf{1}_{A^{\complement}_{n}} \exp\left(\frac{1}{(np_n)^2}\sum_{i,j,k=1}^{n}\hat{\xi}_{i,j}\hat{\xi}_{i,k}d_{ijk}\right)\right] \le \exp\left\{\frac{2d_\star}{p_n}\exp\left\{2K^2 d_\star\right\}\right\}, \tag{3.23}$$

which gives (3.16). □

We are now ready for

Proposition 3.3. *If* (2.1) *holds, then for all* $T > 0$,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\sup_{t\in[0,T]}\left|\theta_t^{i,n} - \bar{\theta}_t^{i,n}\right|^2 = 0, \quad \mathbf{P}\otimes\mathbb{P}\text{-}a.s.. \tag{3.24}$$

Proof. For $i \in \{1, \dots, n\}$, consider

$$\left|\theta_t^{i,n} - \bar{\theta}_t^{i,n}\right|^2 =$$

$$2\int_0^t \left(\theta_s^{i,n} - \bar{\theta}_s^{i,n}\right)\left(F\left(\theta_s^{i,n}\right) - F\left(\bar{\theta}_s^{i,n}\right) + \frac{1}{n}\sum_{j=1}^n \left[\frac{\xi_{i,j}^{(n)}}{p_n}\Gamma\left(\theta_s^{i,n}, \theta_s^{j,n}\right) - \Gamma\left(\bar{\theta}_s^{i,n}, \bar{\theta}_s^{j,n}\right)\right]\right) \mathrm{d}s \leq$$

$$2L_F \int_0^t \left|\theta_s^{i,n} - \bar{\theta}_s^{i,n}\right|^2 \mathrm{d}s + 2L_\Gamma \frac{1}{n}\sum_{j=1}^n \frac{\xi_{i,j}^{(n)}}{p_n}\int_0^t \left(\left|\theta_s^{i,n} - \bar{\theta}_s^{i,n}\right| + \left|\theta_s^{j,n} - \bar{\theta}_s^{j,n}\right|\right)\left|\theta_s^{i,n} - \bar{\theta}_s^{i,n}\right| \mathrm{d}s$$

$$+ 2\int_0^t \left|\frac{1}{n}\sum_{j=1}^n \left(\frac{\xi_{i,j}^{(n)}}{p_n} - 1\right)\Gamma\left(\bar{\theta}_s^{i,n}, \bar{\theta}_s^{j,n}\right)\right| \left|\theta_s^{i,n} - \bar{\theta}_s^{i,n}\right| \mathrm{d}s, \quad (3.25)$$

which gives

$$\left|\theta_t^{i,n} - \bar{\theta}_t^{i,n}\right|^2 \leq$$

$$(2L_F + 1)\int_0^t \left|\theta_s^{i,n} - \bar{\theta}_s^{i,n}\right|^2 \mathrm{d}s + L_\Gamma \frac{1}{n}\sum_{j=1}^n \frac{\xi_{i,j}^{(n)}}{p_n}\int_0^t \left[3\left|\theta_s^{i,n} - \bar{\theta}_s^{i,n}\right|^2 + \left|\theta_s^{j,n} - \bar{\theta}_s^{j,n}\right|^2\right] \mathrm{d}s$$

$$+ \int_0^t \left|\frac{1}{n}\sum_{j=1}^n \left(\frac{\xi_{i,j}^{(n)}}{p_n} - 1\right)\Gamma\left(\bar{\theta}_s^{i,n}, \bar{\theta}_s^{j,n}\right)\right|^2 \mathrm{d}s. \quad (3.26)$$

Summing over i and dividing by n, one obtains

$$\frac{1}{n}\sum_{i=1}^n \left|\theta_t^{i,n} - \bar{\theta}_t^{i,n}\right|^2 \leq$$

$$\leq \left(2L_F + 1 + L_\Gamma \sup_{i=1,\dots,n}\sum_{j=1}^n \frac{3\xi_{i,j}^{(n)} + \xi_{j,i}^{(n)}}{np_n}\right)\int_0^t \frac{1}{n}\sum_{i=1}^n \left|\theta_s^{i,n} - \bar{\theta}_s^{i,n}\right|^2 \mathrm{d}s$$

$$+ \frac{1}{n}\sum_{i=1}^n \int_0^t \left|\frac{1}{n}\sum_{j=1}^n \left(\frac{\xi_{i,j}^{(n)}}{p_n} - 1\right)\Gamma\left(\bar{\theta}_s^{i,n}, \bar{\theta}_s^{j,n}\right)\right|^2 \mathrm{d}s. \quad (3.27)$$

In order to bound $\sum_{j=1}^n \frac{3\xi_{i,j}^{(n)} + \xi_{j,i}^{(n)}}{np_n}$ for all $i = 1, \dots, n$, we choose $K > K_C$ and use Lemma 3.1 to obtain that

$$\sup_{i=1,\dots,n}\sum_{j=1}^n \frac{3\xi_{i,j}^{(n)} + \xi_{j,i}^{(n)}}{np_n} \leq 4 + 4K\,, \quad (3.28)$$

$\mathbb{P}(\,\mathrm{d}\xi)$-a.s.. The application of Gronwall lemma to

$$S_n(t) = \frac{1}{n}\sum_{i=1}^n \left|\theta_t^{i,n} - \bar{\theta}_t^{i,n}\right|^2, \quad (3.29)$$

leads to

$$S_n(t) \leq \int_0^t \exp\left\{G(t-s)\right\}\left(\frac{1}{n}\sum_{i=1}^n \Delta_i(s)\right) \mathrm{d}s, \quad (3.30)$$

with $G = 2L_F + 1 + (4+4K)L_\Gamma > 0$ and $\Delta_i(s)$ defined in (3.5). Therefore

$$\sup_{t\in[0,T]} S_n(t) \le \exp\{GT\} \int_0^T \frac{1}{n}\sum_{i=1}^n \Delta_i(s)\,\mathrm{d}s. \tag{3.31}$$

The last estimate is true for all realizations of the Brownian motions. Taking the limit for n which tends to ∞ and integrating the RHS of (3.31), first with respect to $\mathbb{P}$ (recall Lemma 3.2), completes the proof of Proposition 3.3. □

Proof of Theorem 2.1. Since we already know that $\bar{\mu}^n_\cdot$ converges $\mathbf{P}$-a.s. to $\mu_\cdot$ in $C^0([0,T];\mathcal{P}(\mathbb{R}))$ (see [14, Theorem 1.6]), it suffices to show that

$$\lim_{n\to\infty} \sup_{0\le t\le T} \mathrm{d}_{\mathrm{bL}}\left(\mu^n_t, \bar{\mu}^n_t\right) = 0, \quad \mathbf{P}\otimes\mathbb{P}\text{-a.s.}. \tag{3.32}$$

For every $f:\mathbb{R}\to[0,1]$ 1-Lipschitz function, we have

$$\left|\int_{\mathbb{R}} f(\theta)\left(\mu^n_t - \bar{\mu}^n_t\right)(\mathrm{d}\theta)\right| = \left|\tfrac{1}{n}\sum_{i=1}^n f(\theta^{i,n}_t) - f(\bar{\theta}^{i,n}_t)\right| \le \tfrac{1}{n}\sum_{i=1}^n \left|\theta^{i,n}_t - \bar{\theta}^{i,n}_t\right|. \tag{3.33}$$

In particular,

$$\sup_{0\le t\le T} \mathrm{d}_{\mathrm{bL}}\left(\mu^n_t, \bar{\mu}^n_t\right) \le \sqrt{\frac{1}{n}\sum_{i=1}^n \sup_{0\le t\le T}\left|\theta^{i,n}_t - \bar{\theta}^{i,n}_t\right|^2}. \tag{3.34}$$

The proof follows from Proposition 3.3. □

4. Large Deviation results: proofs.

The proof of Theorem 2.2 relies on two results that contain most of the work. We first prove Theorem 2.2 assuming these two results, and prove them right after.

Proof of Theorem 2.2. Observe that we can write (1.2) as

$$\mathrm{d}\theta^{i,n}_t = F\left(\theta^{i,n}_t\right)\mathrm{d}t + \frac{1}{n}\sum_{j=1}^n \Gamma\left(\theta^{i,n}_t, \theta^{j,n}_t\right)\mathrm{d}t + \sigma\left(\theta^{i,n}_t\right) c_i\left(\theta^n_t\right)\mathrm{d}t + \sigma\left(\theta^{i,n}_t\right)\mathrm{d}B^i_t, \tag{4.1}$$

with

$$c_i\left(\theta^n_t\right) := \frac{1}{n\sigma\left(\theta^{i,n}_t\right)}\sum_{j=1}^n \left(\frac{\xi^{(n)}_{i,j}}{p_n} - 1\right)\Gamma\left(\theta^{i,n}_t, \theta^{j,n}_t\right). \tag{4.2}$$

Recall that $\mathbf{P}^\xi_n$, respectively $\mathbf{P}_n$, is the law of the trajectories $\{\theta^{i,n}_t\}_{i=1,\dots,n;\,t\in[0,T]}$, respectively the law of $\{\bar{\theta}^{i,n}_t\}_{i=1,\dots,n;\,t\in[0,T]}$. The Radon-Nikodym derivative $\mathrm{d}\mathbf{P}^\xi_n/\mathrm{d}\mathbf{P}_n$ is $\exp(M^n_T - \langle M^n\rangle_T/2)$ with

$$M^n_T = \sum_{i=1}^n \int_0^T c_i(\theta^n_t)\,\mathrm{d}\theta^{i,n}_t \quad\text{and}\quad \langle M^n\rangle_T = \sum_{i=1}^n \int_0^T c_i^2(\theta^n_t)\,\mathrm{d}t\,. \tag{4.3}$$

The following lemma is given for every realization of $\xi^{(n)}$ and it has a deterministic nature. Recall that $\chi = \mathcal{P}\left(C^0([0,T];\mathbb{R})\right)$ and $\overline{L}_n$ defined in (2.4). Then $\overline{P}_n(\cdot) := \mathbf{P}(\overline{L}_n \in \cdot)$ is the law of the empirical process associated to (1.4) and $P^\xi_n(\cdot) := \mathbf{P}(L_n \in \cdot)$ is the one associated to (1.2), $\overline{P}_n$ and P^ξ_n are probabilities on χ.

Lemma 4.1. *Suppose $\overline{P}_n(\cdot)$ satisfies a LDP on χ with rate function I. If, for every $C \in \mathbb{R}$,*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbf{E}_n \left[\exp\left\{C\langle M^n\rangle_T\right\}\right] = 0, \tag{4.4}$$

then $P_n^\xi(\cdot)$ satisfies a LDP on χ with the same rate function as $\overline{P}_n$.

Since we want the LDP to hold $\mathbb{P}(\,\mathrm{d}\xi)$-a.s., we need to show that condition (4.4) holds in this sense. To this aim, we redefine the sets Ω_n given in (3.10), as

$$\Omega_n := \left\{\xi : \frac{1}{n} \log \mathbf{E}_n \left[\exp\left\{C_n\langle M^n\rangle_T\right\}\right] > \delta_n\right\}, \tag{4.5}$$

where δ_n and $1/C_n$ tend to zero: they have to do so in a slow way and arbitrarily slow will do for us (explicit choices will be given at the end of the proof).

We need also

$$\Omega^* = \left\{\xi : \text{ there exists } n_0 \text{ s.t. } \frac{1}{n} \log \mathbf{E}_n \left[\exp\left\{C_n\langle M^n\rangle_T\right\}\right] \le \delta_n \text{ for every } n \ge n_0\right\}. \tag{4.6}$$

Lemma 4.2. *Assuming* (2.1) *we have that* $\mathbb{P}(\Omega^*) = 1$.

One readily sees that Lemma 4.2 provides the missing ingredient and the proof of Theorem 2.2 is complete. □

Proof of Lemma 4.1. Recall (4.1)-(4.3). We have to show that (2.3) holds. Consider A a measurable set and recall that A° is the interior of A and $\bar{A}$ is its closure.

Let $p, q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Then

$$P_n^\xi(A^\circ) = \mathbf{P}_n^\xi\left(\{\mu_t^n\}_{t\in[0,T]} \in A^\circ\right) = \mathbf{E}_n\left[1_{\left\{\{\mu_t^n\}_{t\in[0,T]} \in A^\circ\right\}} \exp\left\{M_T^n - \tfrac{1}{2}\langle M^n\rangle_T\right\}\right] \tag{4.7}$$

and Hölder inequality gives

$$P_n^\xi(A^\circ) \ge \left(\overline{P}_n(A^\circ)\right)^p \left(\mathbf{E}_n\left[\exp\left\{-\frac{q}{p}M_T^n - \frac{1}{2}\frac{q}{p}\langle M^n\rangle_T\right\}\right]\right)^{-\frac{p}{q}}. \tag{4.8}$$

Now observe that Cauchy-Schwarz inequality together with the fact that an exponential martingale has expectation less or equal to 1 (see [15, Theorem 5.2]) imply

$$\mathbf{E}_n\left[\exp\left\{-\frac{q}{p}M_T^n - \frac{1}{2}\frac{q}{p}\langle M^n\rangle_T\right\}\right] \le \mathbf{E}_n\left[\exp\left\{\left(\tfrac{2q^2}{p^2} + \frac{q}{p}\right)\langle M^n\rangle_T\right\}\right]^{\frac{1}{2}}. \tag{4.9}$$

Hence

$$P_n^\xi(A^\circ) \ge \left(\overline{P}_n(A^\circ)\right)^p \left(\mathbf{E}_n\left[\exp\left\{\left(\tfrac{2q^2}{p^2} - \frac{q}{p}\right)\langle M^n\rangle_T\right\}\right]\right)^{-\frac{p}{2q}}. \tag{4.10}$$

In particular, one obtains

$$\liminf_{n\to\infty} \tfrac{1}{n} \log P_n^\xi(A^\circ) \ge -p \inf_{x\in A^\circ} I(x) - \tfrac{p}{2q} \liminf_{n\to\infty} \frac{1}{n} \log \mathbf{E}_n\left[\exp\left\{C\langle M^n\rangle_T\right\}\right], \tag{4.11}$$

with $C = \left(\frac{2q^2}{p^2} - \frac{q}{p}\right)$. By hypothesis the second term on the right is zero and since

$$\liminf_{n\to\infty} \tfrac{1}{n} \log P_n^\xi(A^\circ) \ge -p \inf_{x\in A^\circ} I(x), \tag{4.12}$$

is true for all $p > 1$, the lower bound in (2.3) is established.

The upper bound is almost the same: let $p, q > 1$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. Similarly

$$P_n^\xi\left(\bar{A}\right) \le \left(\overline{P}_n\left(\bar{A}\right)\right)^{\frac{1}{p}} \left(\mathbf{E}_n\left[\exp\left\{qM_T^n - \frac{1}{2}q\langle M^n\rangle_T\right\}\right]\right)^{\frac{1}{q}}, \tag{4.13}$$

and, using the properties of exponential martingales as in (4.9), one gets

$$P_n^\xi\left(\bar{A}\right) \le \left(\overline{P}_n\left(\bar{A}\right)\right)^{\frac{1}{p}} \left(\mathbf{E}_n\left[\exp\left\{(2q^2 - q)\langle M^n\rangle_T\right\}\right]\right)^{\frac{1}{2q}}. \tag{4.14}$$

Finally, the desired inequality reads

$$\limsup_{n\to\infty} \tfrac{1}{n} \log P_n^\xi\left(\bar{A}\right) \le -\frac{1}{p} \inf_{x\in\bar{A}} I(x) + \frac{1}{2p} \liminf_{n\to\infty} \frac{1}{n} \log \mathbf{E}_n\left[\exp\left\{C\langle M^n\rangle_T\right\}\right], \tag{4.15}$$

with $C = 2q^2 - q$. And we conclude as before. □

Proof of Lemma 4.2. We want to show that

$$\sum_{n\in\mathbb{N}} \mathbb{P}\left(\Omega_n\right) < \infty. \tag{4.16}$$

As in the proof of Lemma 3.2, let $K > 2$ and consider the events A_n defined in (3.12),

$$A_n = \bigcup_{i=1}^n A_{n,i} \quad \text{with} \quad A_{n,i} = \left\{\xi^{(n)} : \sum_j \left|\frac{\xi_{i,j}^{(n)}}{p_n} - 1\right| > Kn.\right\}.$$

Following the proof of Lemma 3.1, we use $\mathbb{P}\left(\Omega_n\right) \le \mathbb{P}\left(\Omega_n \cap A_n^{\complement}\right) + \mathbb{P}\left(A_n\right)$ and (3.14), i.e. $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) < \infty$, so that one is left with proving that $\sum_{n\in\mathbb{N}} \mathbb{P}\left(\Omega_n \cap A_n^{\complement}\right) < \infty$.

By Markov's inequality applied to $\mathbb{P}\left(\cdot\,|A_n^{\complement}\right)$ we see that

$$\mathbb{P}(\Omega_n \cap A_n^{\complement}) \le \exp\left(-n\delta_n + \log \mathbf{E}\mathbb{E}[\mathbf{1}_{A_n^{\complement}} \exp(C_n\langle M^n\rangle_T)]\right). \tag{4.17}$$

so it suffices to show that

$$\log \mathbf{E}\mathbb{E}[\mathbf{1}_{A_n^{\complement}} \exp(C_n\langle M^n\rangle_T)] \;=\; o(n\delta_n)\,. \tag{4.18}$$

To lighten the notation we go back to using the centered random variables $\hat{\xi}_{i,j} := \xi_{i,j}^{(n)} - p_n$ (cf. (3.8)). With these notations, $\langle M^n\rangle_T$ can be rewritten as

$$\langle M^n\rangle_T \;=\; \frac{1}{(p_n n)^2} \sum_{i,j,k=1}^n \hat{\xi}_{i,j}\hat{\xi}_{i,k}c_{ijk}, \tag{4.19}$$

where

$$c_{ijk} \;=\; \int_0^T \frac{1}{\sigma^2\left(\theta_t^{i,n}\right)} \Gamma\left(\theta_t^{i,n}, \theta_t^{j,n}\right) \Gamma\left(\theta_t^{i,n}, \theta_t^{k,n}\right) \mathrm{d}t. \tag{4.20}$$

Observe that $|c_{ijk}| \le c_\star$ given the boundness of Γ and the conditions on σ.

The estimation of (4.18) is exactly the same as in (3.16), where d_{ijk} are replaced by $C_n c_{ijk}$ (and $d_\star$ by $C_n c_\star$). Following the same strategy, we get

$$\prod_i \mathbb{E}\left[\mathbf{1}_{A_{n,i}^{\complement}} \exp\left(\frac{C_n}{(np_n)^2} \sum_{j,k} \hat{\xi}_{i,j}\hat{\xi}_{i,k}c_{ijk}\right)\right] \le \left(1 + \frac{2C_n}{np_n} \exp\left\{2C_n K^2 c_\star\right\}\right)^n. \tag{4.21}$$

Therefore

$$\log \mathbf{E}\mathbb{E}\left[\mathbf{1}_{A_n^{\complement}} \exp\left(C_n \langle M^n \rangle_T\right)\right] \leq \frac{2C_n}{p_n} \exp\left\{2C_n K^2 c_\star\right\}. \tag{4.22}$$

Which gives (4.18) when $C_n = o(\log(np_n))$ and $\frac{1}{\delta_n} = o\left(\frac{np_n}{\exp\{c\, C_n\}}\right)$ with $c > 2K^2 c_\star$: choose, for example, $C_n = \sqrt{\log(np_n)}$ and $\delta_n = \frac{1}{\sqrt{np_n}}$. □

Acknowledgments

G.G. thanks Amir Dembo for an insightful discussion. F.C. acknowledges the support from the European Unions Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 665850. H.D. acknowledges support of Sorbonne Paris Cité, in the framework of the "Investissements d'Avenir", convention ANR-11-IDEX-0005, and of the People Programme (Marie Curie Actions) of the European Unions Seventh Framework Programme (FP7/2007-2013) under REA grant agreement n. PCOFUND-GA-2013-609102, through the PRESTIGE programme coordinated by Campus France. G.G. acknowledges the support of grant ANR-15-CE40-0020.

References

[1] S. Bhamidi, A. Budhiraja, and R. Wu, *Weakly Interacting Particle Systems on Inhomogeneous Random Graphs*, arXiv:1612.00801

[2] S. Boucheron, G. Lugosi and P. Massart, *Concentration Inequalities: A Nonasymptotic Theory of Independence*, OUP Oxford (2013).

[3] J. von Brecht, T. Kolokolnikov, A. L. Bertozzi and H. Sun, *Swarming on Random Graphs*, J. Stat. Phys. **151** (2013), 150-173.

[4] P. Dai Pra and F. den Hollander, *McKean-Vlasov Limit for Interacting Random Processes in Random Media*, J. Stat. Phys. **84** (1996), 735-772.

[5] A. Budhiraja, P. Dupuis and M. Fischer, *Large Deviation Properties of Weakly Interacting Processes via Weak Convergence Methods*, Ann. Probab. **40**, no. 1 (2012), 74-102.

[6] H. Chiba and G. S. Medvedev, *The Mean Field Analysis for the Kuramoto Model on Graphs I. The Mean Field Equation and Transition Point Formulas*, arXiv:1612.06493

[7] D. A. Dawson and J. Gärtner, *Large Deviations from the Mckean-Vlasov Limit for Weakly Interacting Diffusions*, Stochastics **20** (1987), 247-308.

[8] S. Delattre, G. Giacomin and E. Luçon, *A note on dynamical models on random graphs and Fokker-Planck equations*, J. Stat. Phys. **165** (2016), no. 4, 785-798.

[9] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Stochastic Modelling and Applied Probability **38**, Springer Berlin Heidelberg (2010).

[10] R. L. Dobrushin, *Vlasov equations*, Func. Anal. Appl. **13** (1979), 115-123.

[11] P. Dupuis and R. S. Ellis, *A Weak Convergence Approach to the Theory of Large Deviations*, John Wiley & Sons, (2011).

[12] J. Feng and T. G. Kurtz, *Large Deviations for Stochastic Processes*, Mathematical Surveys and Monographs **131**, American Mathematical Society (2006).

[13] F. den Hollander, *Large Deviations*, Fields Institute Monographs, American Mathematical Society (2000).

[14] J. Gärtner, *On McKean-Vlasov limit for interacting diffusions*, Math. Nachr. **137** (1988), 197-248.

[15] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, North-Holland Mathematical Library, 2nd Edition, (1989).

[16] P.-E. Jabin and Z. Wang, *Quantitative estimates of propagation of chaos for stochastic systems with $W^{-1,\infty}$ kernels*, Invent. Math. **214** (2018), 523-591.

[17] E. Luçon, *Quenched Large Deviations for Interacting Diffusions in Random Media*, J. Stat. Phys. **166** (2017), 1405-1440.

[18] E. Luçon and C. Poquet, *Long time dynamics and disorder-induced traveling waves in the stochastic Kuramoto model*, Ann. Inst. Henri Poincaré Probab. Stat. **53** (2017), 1196-1240.

[19] H. P. McKean, *A Class of Markov Processes Associated with Nonlinear Parabolic Equations*, PNAS **56** (1966), 1907-1911.
[20] G. S. Medvedev, *The continuum limit of the Kuramoto model on sparse random graphs*, arXiv:1802.03787
[21] S. Méléard, *Asymptotic Behaviour of Some Interacting Particle Systems; McKean-Vlasov and Boltzmann Models*, Probabilistic Models for Nonlinear Partial Differential Equations, Lecture Notes in Mathematics **1627**, Springer Berlin Heidelberg (1996), 42-95.
[22] H. Neunzert, *Mathematical investigations on particle-in-cell methods*, Fluid Dynamics Transactions **9** (1978), 229-254.
[23] K. Oelschläger, *A martingale approach to the law of large numbers for weakly interacting stochastic processes*, Ann. Probab. **12** (1984), 458-479.
[24] G. Reis and R. I. Oliveira, *Interacting Diffusions on Random Graphs with Diverging Degrees: Hydrodynamics and Large Deviations*, https://arxiv.org/abs/1807.06898 (2018).
[25] F. A. Rodrigues, T. K. DM. Peron, P. Ji and J. Kurths, *The Kuramoto model in complex networks*, Physics Reports **610** (2016), 1- 98.
[26] A.-S. Sznitman, *Topics in Propagation of Chaos*, Ecole dEté de Probabilités de Saint-Flour XIX, Springer Lectures Notes in Mathematics **1464** (1991), 165-251.
[27] H. Tanaka, *Limit Theorems for Certain Diffusion Processes with Interaction*, In North-Holland Mathematical Library, Stochastic Analysis. Elsevier **32** (1984), 469-88.

Université Paris Diderot, Sorbonne Paris Cité, Laboratoire de Probabilités Statistique et Modélisation, UMR 8001, F- 75205 Paris, France

CNRS, Université Paris Diderot and Sorbonne Université, Institut de Mathématiques de Jussieu - Paris Rive Gauche, F- 75205 Paris, France

Université Paris Diderot, Sorbonne Paris Cité, Laboratoire de Probabilités Statistique et Modélisation, UMR 8001, F- 75205 Paris, France